CLAIMS

We claim:

1. A method executed by a computer under the control of a program, said computer including a memory for storing said program, said method comprising the steps of:

(A) receiving a protein backbone structure with variable residue positions;

(B) establishing a group of potential rotamers for each of said variable residue positions, wherein at least one variable residue position has rotamers from at least two different amino acid side chains; and

(C) analyzing the interaction of each of said rotamers with all or part of the remainder of said protein backbone structure to generate a set of optimized protein sequences, wherein said analyzing step includes a Dead-End Elimination (DEE) computation.

2. A method executed by a computer under the control of a program, said computer including a memory for storing said program, said method comprising the steps of:

(A) receiving a protein backbone structure with variable residue positions;

(B) classifying each variable residue position as either a core, surface or boundary residue;

(C) establishing a group of potential rotamers for each of said variable residue positions, wherein at least one variable residue position has rotamers from at least two different amino acid side chains; and

(D) analyzing the interaction of each of said rotamers with all or part of the remainder of said protein to generate a set of optimized protein sequences.

3. A method according to claim 2 wherein said analyzing step comprises a DEE computation.

4. A method according to claim 1 or 2 wherein said set of optimized protein sequences comprises the globally optimal protein sequence.

5. A method according to claim 1 or 3 wherein said DEE computation is selected from the group consisting of original DEE and Goldstein DEE.

6. A method according to claim 1 or 2 wherein said analyzing step includes the use of at least one scoring function.

7. A method according to claim 6 wherein said scoring function is selected from the group consisting of a Van der Waals potential scoring function, a hydrogen bond potential scoring function, an atomic solvation scoring function, an electrostatic scoring function and a secondary structure propensity scoring function.

8. A method according to claim 6 wherein said analyzing step includes the use of at least two scoring functions.

9. A method according to claim 6 wherein said analyzing step includes the use of at least three scoring functions.

10. A method according to claim 6 wherein said analyzing step includes the use of at least four scoring functions.

11. A method according to claim 6 wherein said atomic solvation scoring function includes a scaling factor that compensates for over-counting.

12. A method according to claim 1 or 2 further comprising testing at least one member of said set to produce experimental results.

13. A method according to claim 4 further comprising

(D) generating a rank ordered list of additional optimal sequences from said globally optimal protein sequence.

14. A method according to claim 13 wherein said generating includes the use of a Monte Carlo search.

15. A method according to claim 2 wherein said analyzing step step comprises a Monte Carlo computation.

16. A method according to claim 13 further comprising:

(E) testing some or all of said protein sequences from said ordered list to produce potential energy test results.

17. A method according to claim 16 further comprising:

(F) analyzing the correspondence between said potential energy test results and theoretical potential energy data.

18. A method according to claim 1 or 2 further comprising altering at least one supersecondary structure parameter value of said protein backbone structure prior to establishing said potential rotamer group.

19. An optimized protein sequence generated by the method of claim 1 or 2.

20. A nucleic acid sequence encoding a protein sequence according to claim 19.

21. An expression vector comprising the nucleic acid of claim 20.

22. A host cell comprising the nucleic acid of claim 20.

23. A protein having a sequence that is at least about 5% different from a known protein sequence and is at least 20% more stable than the known protein sequence.

24. A computer readable memory to direct a computer to function in a specified manner, comprising:

a side chain module to correlate a group of potential rotamers for residue positions of a protein backbone model;

a ranking module to analyze the interaction of each of said rotamers with all or part of the remainder of said protein to generate a set of optimized protein sequences.

25. A computer readable memory according to claim 24 wherein said ranking module includes a van der Waals scoring function component.

26. A computer readable memory according to claim 24 wherein said ranking module includes an atomic solvation scoring function component.

27. A computer readable memory according to claim 24 wherein said ranking module includes a hydrogen bond scoring function component.

28. A computer readable memory according to claim 24 wherein said ranking module includes a secondary structure scoring function component.

29. A computer readable memory according to claim 24 further comprising

an assessment module to assess the correspondence between potential energy test results and theoretical potential energy data.